

### REMARKS

This Amendment After Final is in response to the Final Office Action mailed on December 20, 2007 in which claims 20-21 were allowed, and claims 15-19 and 22-26 were rejected. With this Amendment, claims 15, 17, 22 and 24 are amended. Claims 16, 18-19, 23 and 25-26 are canceled. Claims 15, 17, 22 and 24 are presented for reconsideration and allowance, along with allowed claims 20-21, in view of the following remarks.

#### *Comments on the Examiner's Response to Arguments*

The Examiner's withdrawal of the rejections under 35 U.S.C. § 112 is appreciated. Applicant also thanks the Examiner for the further explanation of the rejections of independent claim 15, and for allowance of claims 20-21. With this Amendment, independent claims 15 and 22 are amended to include limitations, from now canceled dependent claims, which Applicant believes are neither taught nor suggested by Goni. The Examiner's reconsideration of these claims is respectfully requested.

#### *Claim Rejections -35 USC § 102*

In the Final Office Action, claims 15-19 and 22-26 were rejected under 35 U.S.C. § 102(b) as being anticipated by Goni et al., 'ARIES: A lexical platform for engineering Spanish processing tools', Natural Language Engineering 3(4), pg. 317-345, Cambridge University Press, 1997, referred to as "Goni" hereinafter. With this Amendment, independent claims 15 and 22 are amended to include limitations from dependent claims 16 and 18, and from dependent claims 23 and 26, respectively.

As amended, independent claim 15 is directed to a method of annotating verb-clitic form segments in a lexicon, and includes the limitations: "defining, for a segment which is a **clitic pair form segment**, final segment data indicative of whether the segment must appear in a final position of any verb-clitic words formed using the segment; annotating the segment using the defined final segment data; . . . defining, for the segment, word break data indicative of a word break location in the **clitic pair form segment**; and annotating the segment using the word break

data.” (Emphasis added) It is respectfully maintained that the claim limitations relating to the segment being a clitic pair, from now canceled claim 18, are not taught or suggested by Goni.

In rejecting dependent claim 18, the Office Action stated:

Claims 18 and 25: **Goni** discloses a method and tangible computer storage medium as per claim 16 above, wherein:

- i. the segment is a clitic pair form segment, the method further comprising defining, for the segment, word break data indicative of a word break location in the clitic pair form segment (p. 340, Section 5.5.2, *‘If it is not, but the token can be split into a verb form plus some clitic pronouns, ‘)’*; and
- ii. annotating the segment using the word break data (p. 340, Section 5.5.2).

It is respectfully maintained that Goni does not teach or suggest claimed aspects of handling clitic pair form segments. Goni does not teach or suggest defining for the segment, word break data indicative of a word break location in the **clitic pair** form segment, and annotating the segment using the word break data.

In addressing handling of clitic pairs, between page 16, line 16 and page 17, line 10, the present application states:

The segmentation method of the present invention treats some of the full form lexicon entries as strings that are concatenated to create more complex forms. Even though a form created by the attachment of a verb and two clitics is linguistically a combination of three words, for pragmatic reasons some embodiments of the algorithm or method of the present invention handle all verb-clitic forms as pairs of strings. Therefore, it treats combinations of two clitics as one entry in the full form lexicon, and refers to them as “clitic pairs.” Conventionally, linguistic analysis breaks down clitic pairs into their constituent single clitics, which requires complex logic to establish how those pairs combine together and with clitic host variants. One aspect of the invention improves on this by adding the clitic pairs as single entries into the lexicon, and annotating how those pairs combine with clitic host variants. By treating clitic pairs as one single segment or unit in the lexicon, tagging logic required in the prior art is reduced or minimized. Thus, this aspect of the invention minimizes run-time processing requirements without increasing the size of the lexicon by a large number of entries.

Thus, the segmentation algorithm operates only on two segments: an initial segment and a final segment. The first segment is an orthographic/phonological variant of the host form. If the final segment is a clitic

pair, it can be segmented into two pieces for word breaking using information on the word break location stored in the full form lexicon (e.g. *sete* Segs=2 which indicates that the word break location is found after the first two characters “*se*”; and *noslo* Segs=3 which indicates that the word break location is found after the first three characters “*nos*”).

This supports the claim limitation of defining for the segment word break data indicative of a word break location in the **clitic pair** form segment, and annotating the segment using the word break data. In contrast, Goni provides no teaching or suggestion of handling clitic pairs in a lexicon in this manner. The portion of Goni used in rejecting now canceled claim 18 (originally containing these limitations) states:

Once the tagger has been trained, it can be fed with a string or a text to be tagged. First of all, the input is converted into a sequence of tokens which are passed to the Lexicon; if the token is in the lexicon, it is annotated with the tag(s) provided by the Lexicon. If it is not, but the token can be spilt into a verb form plus some clitic pronouns, this decomposition is done, and each of the elements resulting from the splitting is tagged as if it were a token originally obtained from the input text. If the token is neither in the lexicon nor can be split, the suffix table is looked up in the lexicon for a matching; if there is any, the set of tags corresponding to the matched entry is returned. Finally, if all of this fails, a default set of tags is returned.

The output of the tagger could be used as input for a syntactic parser or could be used for compiling statistics on order and relations between different categories of words.

The modifications made to the Xerox Tagger in order to adapt it to Spanish have to do mainly with suffix handling and processing of clitic pronouns (Sanchez and Nieto 1995).

From this passage, it can be seen that Goni addresses the splitting of tokens “into a verb form plus some clitic pronouns.” However, there is no teaching in Goni that the clitic pronouns are clitic pairs (i.e., combinations of two clitics as one entry in the full form lexicon) as described in the present application and claimed in amended claim 15. There is also no teaching in Goni of “defining for the segment, word break data indicative of a word break location in the **clitic pair** form segment, and annotating the segment using the word break data.” Absent a teaching or

suggestion of these claim limitations, amended claim 15 is believed to be in condition for allowance, along with dependent claim 17.

As amended, independent claim 22 is directed to a computer storage medium have stored thereon computer readable instructions for performing steps of a method of annotating verb-clitic form segments in a lexicon. As recited in claim 22, a form segment is a clitic host variant. As amended, claim 22 includes the limitations, from now canceled claim 26, of “defining, for the segment, verification data indicative of whether the clitic host variant **must be combined with a clitic segment to form a valid word**; and annotating the segment using the verification data.” (Emphasis added) It is respectfully maintained that these claim limitations are not taught or suggested by Goni.

In rejecting dependent claim 26, the Office Action stated:

Claims 19 and 26: **Goni** discloses a method and tangible computer storage medium as per claim 16 above, wherein:

- i. the segment is a clitic host variant, the method further comprising defining, for the segment, verification data indicative of whether the clitic host variant must be combined with a clitic segment to form a valid word (p. 325, Table 1 discloses a labeling convention used in modeling verbs in which numbers are assigned to indicate variant forms necessary for each situation); and
- ii. annotating the segment using the verification data (p. 340, Section 5.5.2).

It is respectfully maintained, however, that Goni does not teach the aforementioned limitations.

For instance, in the above-cited portion at page 325, Goni states:

In the conjugation table (Table 1), the *stem\_type* (stt) values of the grammatical features person-number (heading row) and tense-mood (heading column) are displayed in boldface. For example, **sing\_1** means first person, singular number; while **pres.ind** means present tense, indicative mood.

Each of the 49 entries<sup>9</sup> is represented by a numeric code<sup>10</sup>, and the additional value 100 is used as a shorthand for the disjunction of all of them (used for regular verbs; see the entry example in figure 2). The contextual feature *stem\_type* (stt) is used to identify the verb stem and ending corresponding to each form, and the contextual feature *suffix\_type* (sut) distinguishes among several allomorphs of the inflectional morpheme by means of a set of values (Table 2).

While this may represent a labeling convention as asserted in the Office Action, Goni does not actually teach or suggest defining, for the segment, verification data indicative of whether the clitic host variant must be combined with a clitic segment to form a valid word. Identification of verb stems and endings corresponding to each form is not, by itself, a teaching of defining whether a clitic host variant must be combined with a clitic segment to form a valid word, nor of the corresponding annotation of the segment with verification data. Consequently, it is respectfully submitted that independent claim 22 and dependent claim 24 are in condition for allowance over Goni. Reconsideration and allowance of all pending claims are respectfully requested.

The Director is authorized to charge any fee deficiency required by this paper or credit any overpayment to Deposit Account No. 23-1123.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By:                     /John D. Veldhuis-Kroeze/                      
John D. Veldhuis-Kroeze, Reg. No. 38,354  
900 Second Avenue South, Suite 1400  
Minneapolis, Minnesota 55402-3244  
Phone: (612) 334-3222  
Fax: (612) 334-3312

JVK/jmt